基于 CNN-LSTM 的 QAR 数据特征提取与预测 *

张　鹏[1]，杨　涛[2]，刘亚楠[2]，樊志勇[1]，段照斌[1]

(1. 中国民航大学 工程技术训练中心，天津 300300; 2. 中国民航大学 电子信息与自动化学院，天津 300300)

摘　要：针对传统数据驱动的故障诊断方法难以从 QAR 数据中提取有效特征的问题，提出一种融合卷积神经网络(convolutional neural network，CNN)与长短时记忆网络(long short-term memory，LSTM)的双通道融合模型 CNN-LSTM。CNN 与 LSTM 分别作为两个通道，通过注意力机制(attention)融合，从而使模型能同时表达数据在空间维度和时间维度上的特征，并以时间序列预测的方式验证融合模型特征提取的有效性。实验结果表明，双通道融合模型与单一的 CNN、LSTM 相比，能够更有效地提取数据特征，模型单步预测与多步预测误差平均降低 35.3%。为基于 QAR 数据的故障诊断提供一种新的研究思路。

关键词：深度学习；CNN；LSTM；特征提取；时间序列预测

中图分类号：TP391　　doi: 10.3969/j.issn.1001-3695.2018.04.0214

Feature extraction and prediction of qar data based on CNN-LSTM

Zhang Peng[1], Yang Tao[2], Liu Yanan[2], Fan Zhiyong[1], Duan Zhaobin[1]

(1. *Engineering Training Center*, *Civil Aviation University of China*, *Tianjin* 300300, *China*; 2. *College of Electronic Information & Automation*, *Civil Aviation University of China*, *Tianjin* 300300, *China*)

Abstract: Aiming at the problem that it is difficult for the traditional fault diagnosis method to extract effective features from QAR data, this paper proposed a dual-channel fusion model, CNN-LSTM, which combines the Convolutional Neural Network(CNN) and the Long Short-Term Memory(LSTM) . Respectively as two channels, CNN and LSTM fused through the attention mechanism to make the model able to simultaneously express the features of the data in both space dimension and time dimension and verified the validity of the feature extraction of the fusion model through time series prediction. Results of the experiment show that when compared with single CNN or LSTM, the dual-channel fusion model can extract data features more effectively, make the errors of both single-step prediction and multi-step prediction reduce by an average of 35.3%, and provide a new research idea for fault diagnosis based on QAR data.

Key words: deep learning; CNN; LSTM; feature extraction; time series prediction

0　引言

QAR 用于记录飞机在飞行过程中的各项参数，连续完整地反映了飞机系统在运行中的实际状态或失效的征兆信号[1]，对飞行安全性起着至关重要的作用。由于 QAR 数据所记录的多是来自于飞机上各传感器测量的直接参数，无法直观的得出各参数间的关联性，难以直接对飞行品质进行评价。从 QAR 数据中提取反应系统运行状态的特征向量，从而发掘飞机各系统之间的潜在关联性，并将其应用于各系统的故障诊断当中，是该领域一大挑战性问题。

以飞机操纵系统中的动力控制组件(power control unit, PCU)故障诊断为例，故障现象有 PCU 不跟随、PCU 锁定阀门错位、PCU 旁路阀门错位、PCU 作动可靠性测试未通过；涉及的交联系统包括：电源系统、液压系统、传感系统等。在实际排故中，需要维护人员花费大量时间按照复杂的信号分析逻辑对每一种可能产生故障的原因进行排除，维护人员不仅需要掌握本系统的工作原理，还需要了解交联系统的工作原理。如果将 PCU 的故障诊断抽象为一个分类问题，从监控 PCU 的相关 QAR 数据中提取特征，建立特征与故障类别的对应关系，并进行智能故障诊断，对提高排故效率、保障航班正常运行与飞行安全具有重要意义。解决上述问题的核心之一在于如何从 QAR 数据中有效提取完备的特征向量。

收稿日期：2018-04-02；**修回日期：**2018-05-16　**基金项目：**国家自然基金民航联合研究基金重点支持项目（U1533201）；中央高校基本科研业务费专项资助项目（3122016D006）

作者简介：张鹏（1963-），男，北京人，教授，主要研究方向为航空机载设备故障诊断等领域；杨涛（1994-），男（通信作者），湖南邵阳人，硕士研究生，主要研究方向为数据挖掘与故障诊断（tyang_yjs16@163.com）；刘亚楠（1993-），女，河北沧州人，硕士研究生，主要研究方向为电学层析成像；樊志勇（1979-），男，山西人，副教授，主要研究方向为机载电子设备自动测试技术；段照斌（1989-），男，河南人，助教，主要研究方向为航空电子.

传统数据驱动的故障诊断方法需要基于信号处理理论和统计理论从原始数据中人工提取特征，不足之处在于需要大量工程实践与信号处理技术来提取信号特征[2]，该方式严重依赖于人员的专业知识。同时，QAR 数据具有高维、海量的"大数据"特点，使得传统的特征提取方式难以适用。

Hinton 等人[3]提出的深度学习理论是近年来机器学习领域内的研究热点，旨在研究如何从数据中自动地提取多层特征表示[4]。近年来，深度学习在特征提取与模式识别方面显示出独特的优势与潜力，并逐步应用于复杂工业系统的故障诊断当中[5]。文献[6]中提出一种基于深度置信网络(Deep Belief Nets，DBN)[3]的飞行器故障特征诊断方法，通过 DBN 完成故障特征的提取，并实验证明了 DBN 比传统的故障诊断方法具有更高的鲁棒性和准确性；文献[7]首次将 LSTM[8]网络用于锂电池剩余寿命(RUL)预测当中，准确预测了一个周期内的剩余寿命曲线；文献[9]将自动编码器(Autoencoder)用于磨削系统的故障诊断当中，诊断准确率达到 92.4%，远高于传统 BP 神经网络算法。文献[10]中将卷积神经网络(CNN)[11]用于齿轮箱的故障识别和分类当中，并取得了最佳效果。

上述研究表明，基于深度学习的故障诊断算法比传统方法性能上更加优越，能够摆脱传统算法前期特征提取的复杂工作。目前深度学习在计算机视觉、机器翻译、自然语言处理、自动驾驶等领域取得了巨大的成功，但是在故障诊断领域的研究还处于起步阶段，随着深度学习理论的不断完善，其在解决复杂工业系统的故障方面具有很大的潜力[5]。

本文在将研究深度学习理论的基础之上，提出一种双通道融合深度学习模型 CNN-LSTM，用于 QAR 数据特征提取，并以 QAR 数据预测的方式验证所提特征的完备性。在未使用预测参数（俯仰角）历史值的情况下，预测飞机俯仰角在 500 个点内的变化，结果表明，模型特征表达性能更优，单步预测与多步预测精度均高于 CNN、LSTM。

1　深度学习理论

深度学习的概念起源于人工神经网络的研究，有多个隐含层的多层感知器是深度学习模型的显著特征[5]。相对于浅层学习算法，前者具有更好逼近复杂函数的能力，由于具有多隐层结构，能够实现数据特征的逐层转换，保证最有效地信息提取与特征表达[12-13]。深度学习是一种数据驱动的算法，无需建立系统的准确物理模型，只需要收集系统运行的历史数据，即可获得系统的最优特征表示，从而完成故障诊断、故障分类、故障预测等任务。本文主要研究 CNN 与 LSTM 两种算法。

1.1　卷积神经网络（CNN）

CNN 是由 LeCun[11]提出，是一种前馈神经网络。典型的卷积神经网络 LeNet-5 模型如图 1 所示。其中包括输入层（input）、卷积层（C）、池化层（S）、全连接层（F）以及输出层（output）。

CNN 的本质在于构建多个能够提取数据特征的滤波器，通过对输入数据进行逐层卷积和池化操作来提取数据之间隐藏的拓扑结构特征。随着层数的不断增加，提取的特征越来越抽象，最后将这些抽象特征通过全连接层汇合，并通过 softmax 或 sigmoid 激活函数解决分类问题和回归问题。CNN 的特点之一在于可提取输入数据的局部特征，并逐层组合抽象生成高层特征，有效实现故障诊断与识别。

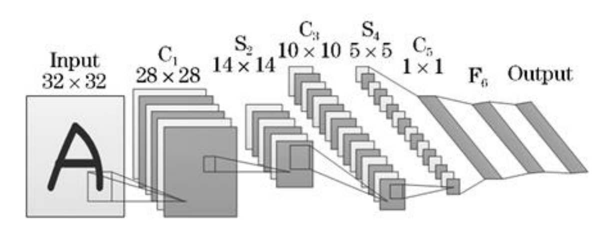


图 1　LeNet-5 模型

1.2　长短时记忆网络（LSTM）

LSTM 是一种时间递归的神经网络，非常适合处理时间序列数据。相较与传统的递归神经网络（RNN），LSTM 通过引入了记忆单元，解决了 RNN 在实际应用中长期依赖的问题。

LSTM 由输入层、隐含层和输出层构成，其结构如图 2 所示。其中，$g(t)$ 表示输入单元，$h(t)$ 表示状态输出单元，M 表示记忆单元，$i(t)$、$o(t)$、$f(t)$ 分别表示输入门限、输出门限以及遗忘门限。

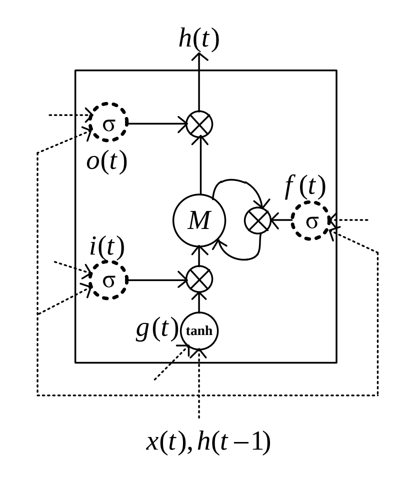


图 2　LSTM 结构

由图 2 可以看出，M 的读、写和遗忘操作由三个门限单元来控制。设输入时间序列为 x，t 为当前时刻，则各单元的状态可用如下公式表示。

输入单元：

$$g(t)=\tanh(W_{xg}g(t-1)+W_{hg}h(t-1)+b_g) \tag{1}$$

门控单元：

$$i(t)=\sigma(W_{xi}i(t-1)+W_{hi}h(t-1)+b_i) \tag{2}$$

$$f(t)=\sigma(W_{xf}i(t-1)+W_{hf}h(t-1)+b_f) \tag{3}$$

$$o(t)=\sigma(W_{xo}i(t-1)+W_{ho}h(t-1)+b_o) \tag{4}$$

记忆单元：

$$M(t)=f(t)M(t-1)+i(t)g(t) \tag{5}$$

状态输出单元：

$$h(t)=o(t)\tanh(M(t)) \tag{6}$$

其中，σ 表示 sigmoid 激活函数。

2　双通道融合模型

CNN 与 LSTM 都是深度学习的主流算法，相对而言，CNN 更适合进行空间扩展，提取数据的局部特征，并组合抽象成高层特征。而 LSTM 更适合进行时间扩展，具有长期记忆功能，更适合处理时间序列。在 QAR 数据的特征提取中，不仅需要考虑不同参数之间的空间联系，还需要考虑数据在时间维度上的变化，因此本文将 CNN 与 LSTM 两种模型以 Attention[14]的方式相结合，提出一种双通道融合模型(CNN-LSTM)，使模型具有时间、空间的特征表达能力，模型结构如图 3 所示。

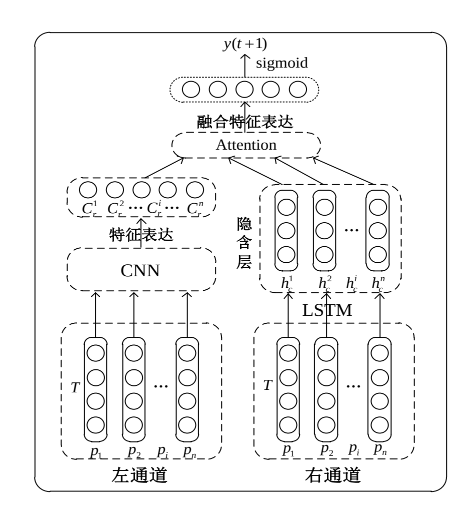


图 3　CNN-LSTM 模型结构

其中，p_i 为输入样本中的第 i 个参数，T 为样本的时间长度，C_r^i 为 CNN 网络提取的第 i 个特征，h_c^i 为 LSTM 第 i 个隐含层的输出。

样本分别从左右两个通道输入网络，其中左通道为 CNN 网络。CNN 网络的最后一层由 n 个神经元输出，得到 $1\times n$ 维特征向量：

$$C_r=[c_r^1,c_r^2,\cdots,c_r^n] \tag{7}$$

右通道为 LSTM 网络，序列长度为 n，LSTM 隐含层输出维度为 m，因此得到 $n\times m$ 维特征向量：

$$H_c=[h_c^1,h_c^2,\cdots,h_c^n] \tag{8}$$

得到样本数据在两个通道的特征表达之后，采用 Attention 机制得到融合特征图，Attention 层结构如图 4 所示：

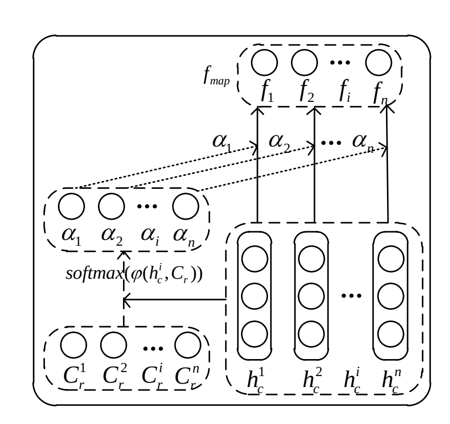


图 4　Attention 层结构

其过程可按如下公式描述：

$$\varphi(h_c^i,C_r)=\tanh(h_c^i\bullet W_a\bullet C_r^T+b_a) \tag{9}$$

$$\alpha_i=\frac{\exp(\varphi(h_c^i,C_r))}{\sum_{j=1}^{n}\exp(\varphi(h_c^j,C_r))} \tag{10}$$

$$f_{map}=\sum_{i=1}^{n}\alpha_i h_c^i \tag{11}$$

其中：W_a 为 $m\times n$ 维的权值矩阵，b_a 为偏置项，均由训练时学习得到。C_r^T 是 C_r 的转置。

式(9)将 CNN 的特征向量与 LSTM 的特征向量融合成一组“权重”，并通过式(10)Softmax 激活函数将“权重”归一化到 0-1 之间。最后通过式(11)，将 LSTM 不同时间点的隐含层输出向量 h_c^i 与对应的“权重” α_i 相乘，并求和，以此作为样本最终的特征表达，即 f_{map}。其意义在于：将 CNN 组合抽象局部特征的能力通过注意力机制的方式来调节 LSTM 时间维度的特征表达，使其特征在某些区域加强，在某些区域减弱，从而使模型具有时间、空间的特征表达能力。

3　算例分析

3.1　数据集构造

本文以飞行控制系统中的俯仰通道为研究对象，从相关参数中提取特征，来预测俯仰角的变化情况。根据飞控系统工作原理[15]，从 QAR 数据选取 22 个相关参数，采样间隔为 1 秒，各参数含义如表 1 所示。

表 1　相关参数表

参数名称	参数注释	单位
ALTITUDE (1013)	高度	FT
COMPUTED AIR SPEED	计算空速	KT
VERTICAL ACCEL	垂直加速度	G
FLIGHT PATH ACCELERATION	飞航加速度	G
ALTITUDE RATE	高度变化率	FT/MIN
WIND SPEED	风速	KT
DRIFT ANGLE	偏流角	DEG
BODY PITCH RATE	机体俯仰速率	DEG/SEC
BODY ROLL RATE	机体滚转速率	DEG/SEC
BODY YAW RATE	机体偏航速率	DEG/SEC
FLIGHT DIRECTOR - PITCH	飞行指引-俯仰角	DEG
FLIGHT DIRECTOR - ROLL	飞行指引-滚转角	DEG
FLIGHT PATH ANGLE	飞行航迹角	DEG
GROUND SPEED	地速	KT
INDICATED AOA	迎角指示	DEG
MACH NUMBER	马赫数	M
PITCH ATTITUDE	俯仰角	DEG
PITCH ATTITUDE RATE	俯仰角速率	DEG/SEC
ROLL ATTITUDE	滚转角	DEG
STABILIZER POS	水平安定面位置	DEGRVDT

TOTAL AIR TEMPERATURE	大气总温	DEG C
VERTICAL SPEED	垂直速度	FT/MIN

其中，俯仰角(PITCH ATTITUDE)为预测参数，其余数据为训练参数。

样本构造方式如图 5 所示。将归一化的数据按照时间进程依次排列，组成二维数据矩阵形式，以滑动窗口截取方式构造训练集，其中预测标签为 N 时刻之后预测参数 y 的值，如图 5 中的 y_1 、 y_2 。考虑到本实验为一个时间序列预测任务，为保证预测参数在时间进程上的连续性，故选取步长 $s=1$；同时，为便于 CNN 做卷积操作，选取与训练参数个数相同的滑动窗口大小，即 $w=21$，故每个样本为一个 21×21 的方阵。同时，为检验模型单步预测与多步预测的性能，分别选取 $N=1,5,10$ 并按上述方式构造样本得到数据集 S_1 、 S_2 、 S_3 。

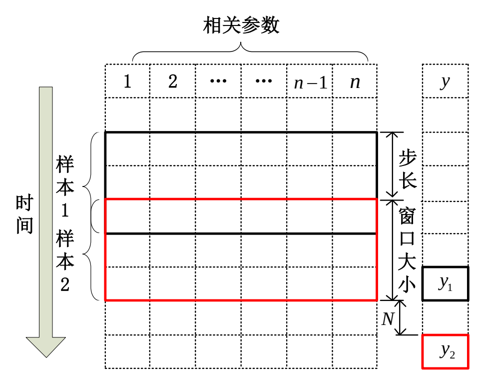


图 5　样本构造方法

S_1 、 S_2 、 S_3 中训练集、验证集、测试集样本划分情况如表 2 所示。

表 2 数据集划分

	训练集	验证集	测试集
S_1	14144	3537	500
S_2	14141	3536	500
S_3	14137	3535	500

3.2　实验与结果

为验证所提双通道融合模型的有效性，在基于 TensorFlow 的 Keras 深度学习平台上实现该模型。并与单一的 LSTM 与 CNN 模型在数据集 S_1 、 S_2 、 S_3 上分别做对比实验。各模型结构参数如图 6 所示。

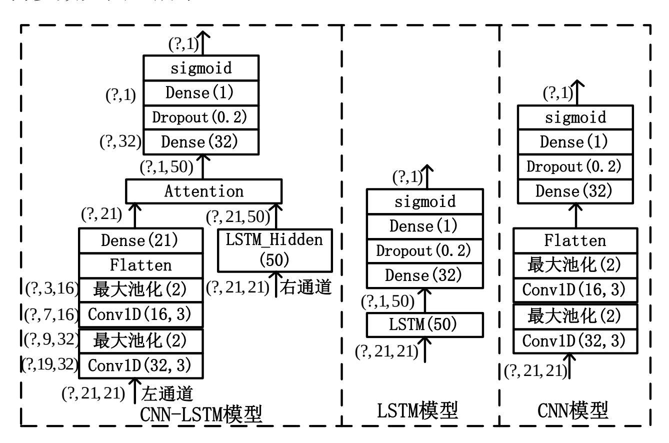


图 6　各模型结构和参数

以数据集 S_1 为例，数据输入形式是一个三维的张量 $(?,21,21)$，? 表示每批次训练送入样本的数量。

在 CNN-LSTM 模型中，左(CNN)通道的数据先经过一维卷积层(Conv1D)，含有 32 个 1×3 的卷积核，激活函数为 Relu 函数，此时张量形状变为 $(?,19,32)$；再经过一个窗口为 2 的最大池化层(Maxpooling1D)，张量维度变为 $(?,9,32)$；然后经过含有 16 个 1×3 卷积核的卷积层，张量维度变为 $(?,7,16)$；再经过窗口为 2 的最大池化层，张量维度变为 $(?,3,16)$；最后由 Flatten 层将数据“压平”，并经过输出维度为 21 的全连接层(Dense)，此时张量维度为 $(?,21)$ 。右(LSTM)通道的数据经一个输出维度为 50 的 LSTM 单元，并取其隐含层输出，得到一个 $(?,21,50)$ 的张量。将两个通道张量由按式(9)实现的 Attention 层融合，张量形状变为 $(?,1,50)$，得到数据的融合特征表达。再通过输出维度为 32 的全连接层，并施加 Dropout 层，防止过拟合，最后由输出维度为 1 的全连接层激活输出，激活函数为 Sigmoid。

LSTM 模型与 CNN-LSTM 右通道类似，输入数据经由一个输出为 50 维的 LSTM 单元，但只取其隐含层最后的输出，输出张量维度为 $(?,1,50)$，之后的处理方式与 CNN-LSTM 模型 Attention 层后的类似。CNN 模型与 CNN-LSTM 左通道类似。

为衡量模型性能，参考相关文献[16]，采用平均绝对误差（MAE）、平均绝对百分比误差（MAPE）、均方根误差（RMSE）、均方根对数误差（RMSLE）作为评价指标，计算公式如式(12)~(15)所示。

$$MAE=\frac{1}{n}\times\sum_{t=1}^{n}\left|A_t-F_t\right| \tag{12}$$

$$MAPE=\frac{1}{n}\times\sum_{t=1}^{n}\left|\frac{A_t-F_t}{A_t}\right|\times100\% \tag{13}$$

$$RMSE=\sqrt{\frac{1}{n}\times\sum_{t=1}^{n}(A_t-F_t)^2} \tag{14}$$

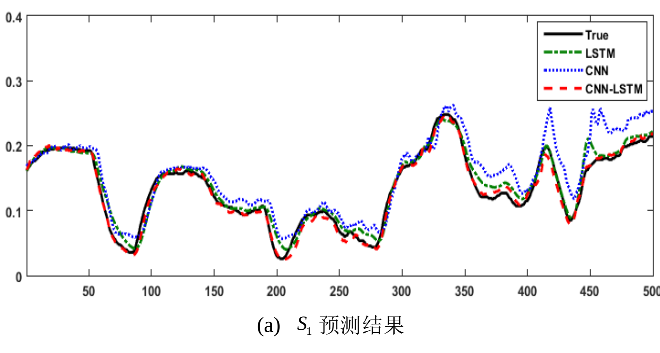
$$RMSLE=\sqrt{\frac{1}{n}\times\sum_{t=1}^{n}(\log(A_t+1)-\log(F_t+1))^2} \tag{15}$$

其中： A_t 为预测值， F_t 为真实值。以上四种指标越小表明预测值越接近真实值，证明模型性能越好、特征表达能力越强。

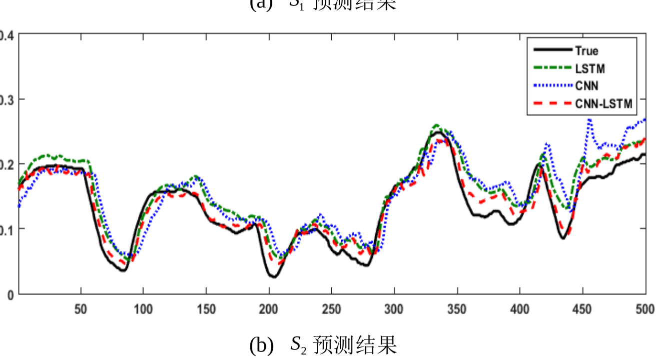
上述三种模型在三份测试集上的预测结果如图 7(a)-(c)所示。误差指标计算如表 3 所示。

比较图 7(a)~(c)的结果可知：在数据集 S_1 上，由于是单步预测，预测时间间隔较短，三种模型都能准确预测出俯仰角真实值，CNN 模型在第 350 个点之后开始偏离真实曲线，但总体趋势与真实曲线一致；在数据集 S_2 上，预测时间间隔增加到 5 步，与 S_1 比较，三种模型均不同程度偏离真实值，其中 CNN-LSTM 模型的偏离程度最小；在数据集 S_3 上，预测时间间隔扩大到 10 步，各模型进一步偏离真实曲线，且整体存在一个滞后效应，在第 450 个点之后，CNN 与 LSTM 均存在偏离真实趋势的震荡，只有 CNN-LSTM 模型与真实曲线趋势保持一致。同时，表 3 中误差计算结果表明：CNN-LSTM 在数据集 S_1 与 S_2 上四种误差指标均明显优于 CNN 于 LSTM；在 S_3 上，

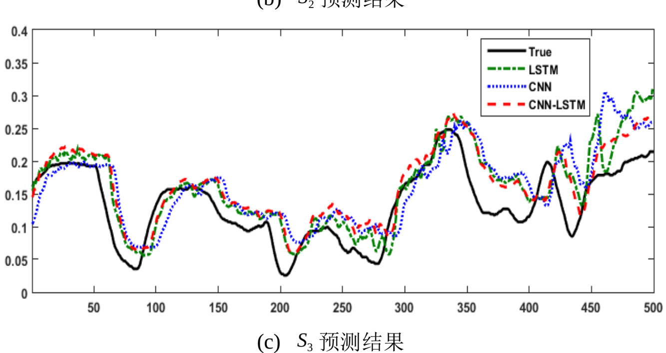
CNN-LSTM 与 LSTM 性能接近，在 MAE 指标上两者持平，在 MAPE 指标上，LSTM 要稍好于 CNN-LSTM；在三份数据集中，LSTM 均好于 CNN，说明 LSTM 比 CNN 更适合处理时间序列数据。结合各项评价指标，在单步预测上：CNN-LSTM 比 CNN 误差平均降低 69.8%，比 LSTM 误差平均降低 38.2%；在五步预测上：CNN-LSTM 比 CNN 误差平均降低 51.0%，比 LSTM 误差平均降低 33.1%；在十步预测上：CNN-LSTM 比 CNN 误差平均降低 19.6%，比 LSTM 误差平均降低 1.5%；综合而言，CNN-LSTM 单步预测与多步预测误差比单一 CNN 与 LSTM 平均降低 35.3%。



(a) S_1 预测结果



(b) S_2 预测结果



(c) S_3 预测结果

图 7 各模型预测结果对比

表 3 误差对比结果

		MAE	MAPE	RMSE	RMSLE
S_1	CNN	0.021	15.6%	0.026	0.023
	LSTM	0.009	9.0%	0.012	0.011
	CNN-LSTM	**0.006**	**6.1%**	**0.007**	**0.006**
S_2	CNN	0.029	21.2%	0.035	0.031
	LSTM	0.021	16.5%	0.025	0.022
	CNN-LSTM	**0.013**	**11.5%**	**0.017**	**0.015**
S_3	CNN	0.041	27.0%	0.051	0.044
	LSTM	**0.033**	**22.1%**	0.042	0.036
	CNN-LSTM	**0.033**	22.4%	**0.040**	**0.035**

上述实验结果表明：所提融合方式能够结合 CNN 与 LSTM 各自的特点，更加完备的从 QAR 数据中提取特征。

4 结束语

QAR 数据的有效分析与利用对提高航班运行效率、保障飞行安全显得尤为重要，受限于 QAR 数据高维、海量的样本特性，传统特征提取方式难以适用，目前对 QAR 数据的分析主要集中在统计分析上。本文提出一种基于深度学习的双通道融合模型 CNN-LST M，模型将 CNN 与 LSTM 以 Attention 的方式结合，使得模型能够更完备地提取 QAR 数据的特征，从而为 QAR 数据提供一种更为有效的利用方法。通过对飞行过程中俯仰角参数的预测表明：CNN 与 LSTM 二者结合起来的融合模型在特征表达上比单一模型更有效，单步预测与多步预测精度更高。同时，本文介绍了一种适合工业大数据的数据集构造方式，通过选取不同的窗口大小、步长、时间间隔以适应不同的实际问题。本文为后续开展基于 QAR 数据的故障诊断等提供了一种新的研究思路。

参考文献：

[1] 耿宏，李萍萍. QAR 数据处理在气动力矩参数辨识中的应用 [J]. 航空电子技术, 2012, 43 (03): 25-29. (Geng Hong, Li Pingping. Application of QAR data processing in aerodynamic moment parameter identification [J]. Avionics Technology, 2012, 43 (03): 25-29.)

[2] 鞠建波，胡胜林，祝超，等. 基于深度学习的装备故障诊断方法 [J]. 电光与控制, 2018, 25 (02): 103-106. (Ju Jianbo, Hu Shenglin, Zhu Chao, *et al.* A deep learning based method for equipment fault diagnosis [J]. Electronics Optics &Control, 2018, 25 (02): 103-106.)

[3] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets [J]. Neural Computation, 2006, 18 (7): 1527-1554.

[4] 张军阳，王慧丽，郭阳，等. 深度学习相关研究综述 [J/OL]. 计算机应用研究, 2018, 35 (7) . [2018-03-21]. http://kns. cnki. net/kcms/detail/51. 1196. tp. 20170818. 1709. 084. html. (Zhang Junyang, Wang Huili, Guo Yang, *et al.* Review of Deep Learning [J/OL] , Application Research of Computers, 2018, 35 (7) . [2018-03-21]. http://kns. cnki. net/kcms/detail/51. 1196. tp. 20170818. 1709. 084. html.)

[5] 任浩，屈剑锋，柴毅，等. 深度学习在故障诊断领域中的研究现状与挑战 [J]. 控制与决策, 2017, 32 (8): 1345-1358. (Reng Hao, Qu Jianfeng, Chai Yi, *et al.* Deep learning for fault diagnosis: The state of the art and challenge [J]. Control and Decision, 2017, 32 (8): 1345-1358.)

[6] Jiang Hongkai, Shao Haidong, Chen Xinxia, *et al.* Aircraft Fault Diagnosis Based on Deep Belief Network [C]// Proc of International Conference on Sensing, Diagnostics, Prognostics, and Control. 2017: 123-127.

[7] Zhang Yongzhi, Xiong Rui, He Hongwen, *et al.* A LSTM-RNN method for the lithuim-ion battery remaining useful life prediction [C]// Proc of Prognostics and System Health Management Conference. 2017: 1-4.

[8] Hochreiter S, Schmidhuber J. Long Short-term memory [J]. Neural

Computation, 1997, 9 (8): 1735-1780.

[9] Qu Xingyu, Zeng Peng, Fu Dongdong, *et al.* Autoencoder-based fault diagnosis for grinding system [C]// Proc of Chinese Control and Decision Conference. 2017: 3867-3872.

[10] Chen Zhiqiang, Li Chuan, Sanchez R V. Gearbox fault identification and classification with convolutional neural networks [J]. Shock & Vibration, 2015, 2015 (2): 1-10.

[11] Lecun Y, Bottou L, Bengio Y, *et al.* Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86 (11): 2278-2324.

[12] Jia Feng, Lei Yaguo, Lin Jing, *et al.* Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data [J]. Mechanical Systems & Signal Processing, 2016, 72-73: 303-315.

[13] Zhang Qingchen, Yang L T, Chen Zhikui. Deep computation model for unsupervised feature learning on big data [J]. IEEE Trans on Services Computing, 2016, 9 (1): 161-171.

[14] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate [J]. Computer Science, 2014.

[15] 徐军. 飞行控制系统：设计、原型系统及半物理仿真实验 [M]. 北京：北京理工大学出版社，2015. (Xu Jun. Flight Control System: Design, Prototype and Semi Physical Simulation Experiment [M]. Beijing: Beijing Institute of Technology Press, 2015.)

[16] 张承钊. 一种金融市场预测的深度学习模型: FEPA 模型 [D]. 成都：电子科技大学，2016. (Zhang Chengzhao. A Deep Learning Model for Financial Market Forecasting: FEPA Model [D]. Chengdu: University of Electronic Science and Technology of China, 2016.)